Eur Biophys J (2002) 31: 217–227
DOI 10.1007/s00249-002-0207-5

ARTICLE

José D. Faraldo-Gómez · Graham R. Smith
Mark S.P. Sansom

Setting up and optimization of membrane protein simulations

Received: 13 September 2001 / Revised: 7 January 2002 / Accepted: 7 January 2002 / Published online: 19 February 2002

Abstract In this paper we describe a method for setting up an atomistic simulation of a membrane protein in a hydrated lipid bilayer and report the effect of differing electrostatic parameters on the drift in the protein structure during the subsequent simulation. The method aims to generate a suitable cavity in the interior of a lipid bilayer, using the solvent-accessible surface of the protein as a template, during the course of a short steered molecular dynamics simulation of a solvated lipid membrane. This is achieved by a two-stage process: firstly, lipid molecules whose headgroups are inside a cylindrical volume equivalent to that defined by the protein surface are removed; then the protein-lipid interface is optimized by applying repulsive forces perpendicular to the protein surface, and of gradually increased magnitude, to the remaining lipid atoms inside the volume occupied by the protein surface until it is emptied. The protein itself may then be inserted. Using the bacterial membrane proteins KcsA and FhuA as test cases, we show how the method achieves the formation of a suitable cavity in the interior of a dimyristoylphosphatidylcholine lipid bilayer without perturbing the configuration of the non-interfacial regions of the previously equilibrated lipid bilayer, even in cases of membrane proteins with irregular geometrical shapes. In addition, we compare subsequent simulations in which the long-range electrostatic interactions are treated via either a cut-off or a particle-mesh Ewald method. The results show that the drift from the initial structure is less in the latter case, especially for KcsA and for the non-core secondary structural elements (i.e. surface loops) of both proteins.

Keywords Membrane proteins · Lipid bilayers · Simulations · Molecular dynamics

J.D. Faraldo-Gómez and G.R. Smith contributed equally to this paper

J.D. Faraldo-Gómez · M.S.P. Sansom (✉)
Laboratory of Molecular Biophysics,
Department of Biochemistry, The Rex Richards Building,
University of Oxford, South Parks Road,
Oxford OX1 3QU, UK
E-mail: mark@biop.ox.ac.uk
Tel.: +44-1865-275371
Fax: +44-1865-275182

G.R. Smith
Biomolecular Modelling Laboratory,
Imperial Cancer Research Fund,
44 Lincoln's Inn Fields, London WC2A 3PX, UK

Introduction

Membrane proteins are of almost self-evident importance in current biology, particularly with respect to discovery of new drug targets (Terstappen and Reggiani 2001). They constitute ca. 30% of known genes (Wallin and von Heijne 1998) and, after a long lag, experimental studies are starting to yield a substantial number of membrane protein structures (see e.g. http://blanco.biomol.uci.edu/membrane_proteins_xtal.html). Molecular dynamics (MD) simulations and related methods play an important role in helping us to fully understand the conformational dynamics of membrane proteins (Forrest and Sansom 2000), and hence the relationship between their structure and biological function, as exemplified by recent studies of, for example, ion channels (Roux et al. 2000; Sansom et al. 2000).

Fully atomistic simulations of membrane proteins must include a lipid bilayer or other membrane mimetic [e.g. a slab of octane (Guidoni et al. 1999) or methane (Govaerts et al. 2001)] in order to model the hydrophobic environment in which these proteins exist. Given the slow dynamics of lipid molecules in fluid-phase bilayers [with a diffusion coefficient $D \approx 10^{-3}$–10^{-4} Å^2 ps^{-1} (Forstner et al. 2000; Gennis 1989; Pastor and Feller 1996)], and the irregular shape of (most) membrane proteins, obtaining a correctly configured initial system is a non-trivial task, and yet the reliability of the subsequent simulation may depend on how carefully this is performed. For example, one wishes to have optimal

packing of the lipid tails around the hydrophobic surface of the integral protein.

In order to build these protein-lipid bilayer systems, two approaches have been reported in the literature. The first (Petrache et al. 2000a; Roux and Woolf 1996; Woolf and Roux 1994, 1996) consists of building a bilayer around the protein lipid by lipid, each individual molecule being selected from a library of lipid conformations; these structures are obtained from simulations of either lipid bilayers or individual lipid molecules in a membrane-like environment. Unfavourable lipid-lipid and protein-lipid contacts are removed during a second stage using a rigid-body conformational search; Tang et al. (1999) subsequently conduct a process in which, by contracting the bilayer in the membrane plane, a range of values of the membrane surface area may be explored. The second approach (Shen et al. 1997; Tieleman and Berendsen 1998) uses a previously equilibrated lipid bilayer, in which a cylindrical hole to accommodate the protein is created by the application of weak repulsive radial forces on the lipid atoms. In both cases the protein-lipid system is energy minimized prior to the MD calculation.

Given the availability of a number of well-characterized lipid bilayer models, of diverse size and composition (Berger et al. 1997; Chiu et al. 2001; Feller and Pastor 1999; Lindahl and Edholm 2000; Schneider and Feller 2001; Tieleman et al. 1997; Tobias et al. 1997; Venable et al. 1993), the second method (i.e. protein insertion into a pre-formed bilayer) appears particularly appealing, since it may not require a long equilibration of the lipid bilayer for each system model after the protein has been inserted, provided that the bilayer was well equilibrated to start with, and that the generation of the cavity is such that only the protein-lipid interface is perturbed significantly. Nevertheless, whilst a cylindrical hole is suitable in the case of simulations of single α-helices or some simple α-helical bundles (Tieleman et al. 1999a), it is problematic in the case of proteins of more complicated cross-sectional geometries.

In this paper we report a method that has been developed from this second approach, but that allows the cavity in the bilayer to have an arbitrary shape, thus generalizing the method to any protein geometry. The perturbation of the lipid bilayer is minimized by conducting the process in several stages, in which the force applied to the lipid atoms is gradually increased until a satisfactory cavity is generated. Having successfully inserted two proteins [KcsA (Doyle et al. 1998) and FhuA (Ferguson et al. 1998; Locher et al. 1998)] in a dimyristoylphosphatidylcholine (DMPC) bilayer by this process, we then perform production simulations of duration 1 ns, which have been used to explore the effects of using cut-offs versus particle-mesh Ewald (PME) algorithms (Darden et al. 1993) in the treatment of the electrostatic interactions. Thus, we hope to arrive at an optimal procedure for simulations of membrane proteins in general.

Methods

General simulation methodology

The simulations presented in this report were conducted using versions 1.6 and 2.0 of the GROMACS (Berendsen et al. 1995) MD simulation package (http://www.gromacs.org). The systems to be simulated (Table 1) were set up using a rectangular box, employing periodic boundary conditions to avoid edge effects. All simulations were conducted at constant temperature (310 K), pressure (1 bar) and number of particles (Berendsen et al. 1984). Numerical integration of the equations of motion used a time step of 2 fs, with atomic coordinates saved every 1 ps for analysis. Solvent (i.e. water and ions), lipid and protein were coupled separately to a temperature bath, with a coupling constant of $\tau_T = 0.1$ ps. Anisotropic pressure coupling was used, with a coupling constant of $\tau_P = 1.0$ ps, so that the dimensions of the simulation box are scaled independently in each direction. Energy minimizations were performed using a steepest descent algorithm. When restraints were required on, for example, protein atoms, a harmonic potential of force constant 10 kJ mol^{-1} $Å^{-2}$ was applied.

The extended-atom GROMOS 87 force field (Hermans et al. 1984) was used, as implemented in GROMACS. Bond distances were constrained using the LINCS algorithm (Hess et al. 1997). The van der Waals interactions were modelled using a 6-12 Lennard-Jones (LJ) potential, cut off at 10 Å, with first- and second-neighbour exclusions and scaled third-neighbour LJ coefficients. In the "CO" simulations (see below), electrostatic interactions were cut off at 18 Å; in the "PME" simulations, a particle-mesh Ewald algorithm (Darden et al. 1993) was used, with a 9 Å cut-off for the direct space calculation; the reciprocal space calculation was performed using a fast Fourier transform (FFT) algorithm.

Setting up the systems

The procedure adopted for insertion of a protein into a lipid bilayer aimed to create a protein surface-adapted conformation of those lipids close to the protein-lipid interface, without causing a major disruption in the configuration of the bilayer in other regions. In essence, the method relies on the use of a protein surface to define an *exclusion region* within the membrane, from which water and lipid molecules are to be expelled by means of a force applied during a few tens of picoseconds of molecular dynamics. A similar procedure, in which a cylindrically shaped cavity was obtained, had previously been developed by Shen et al. (1997) and used in simulations of single transmembrane (TM) helices or regular helix bundles (Forrest et al. 1999; Randa et al. 1999; Tieleman et al.

Table 1 Details of simulation systems

Simulation[a]	Number of lipids	Number of waters	Numbers of ions	Total number of atoms	Box dimensions (Å)
KcsA	245	11,383	$3K^+ + 7Cl^-$	49,125	106×80×80
FhuA	224	17,729	$36Na^+ + 31Cl^-$	70,628	106×80×103

[a]For each system, two production run simulations were performed, one with long-range interactions treated via cut-offs (KcsA-CO and FhuA-CO) and one using PME (KcsA-PME and FhuA-PME)

1999a). However, this latter method was not adequate for more irregular cross-sectional shapes, or for proteins with a significant degree of asymmetry with respect to the membrane plane; the use of the protein surface as a template for generating the cavity inside the lipid bilayer is intended specifically to account for these complexities.

The method consists of the following steps, beginning with two coordinate files, one of the solvated and equilibrated lipid bilayer and the other of the protein to be inserted:

1. Generation of a solvent-accessible protein surface, using the program GRASP (Nicholls et al. 1993), to be used as a template for the exclusion region [the method of Sanner et al. (1996) may also be used]. This surface is generated by rolling a 1.4 Å radius sphere over the van der Waals surface of the protein to define a triangular grid of given resolution (in this case, 0.5 grid-points $Å^{-2}$), which in turn defines the accessible surface. In the case of FhuA, the periplasmic mouth of the protein, which forms a large cavity, had been filled with water molecules prior to building the surface, in order to prevent lipid molecules from remaining in this region during the steered MD run (i.e. cavity optimization).
2. Estimation of the volume of the exclusion region, using the program HOLE (Smart et al. 1993, 1996) to calculate the cross-sectional area profile of the protein surface along the direction normal to the membrane plane; by integrating this profile over the TM region of the protein, the volume to be occupied by the protein in each leaflet of the lipid bilayer was estimated. In this context, it was assumed that the TM region of a membrane protein is defined by the mean distance between the two bands of aromatic residues that are thought to be involved in anchoring the protein within the lipid bilayer.
3. For each leaflet of the bilayer, removal of those lipids whose headgroup P atom is contained within a cylinder situated in the centre of the bilayer, its volume being equal to that calculated in step 2. As a result of this deletion of lipids, atoms mainly from lipid tails (and water molecules) remain inside the exclusion region. The removal of lipids is required to prevent dramatic changes in the area per lipid of the bilayer when the MD is performed. If no lipid molecules were removed at this stage, the forces applied would induce an increased density of lipids in the protein-bilayer interface, which would be minimally accommodated by non-interfacial regions, owing to the slow dynamics of lipid molecules in bilayers ($D \approx 10^{-3}$–10^{-4} $Å^2$ ps^{-1}).
4. Generation of the protein surface-shaped cavity. The core MD program MDRUN from GROMACS was modified to read the vertices and surface-normals from the triangulated molecular surface generated in stage 1. Each atom of the lipid and solvating water is associated with its closest surface vertex by a neighbour search. Consider an atom with position $\boldsymbol{r}$, and let its nearest surface vertex have position $\boldsymbol{s}$ and (outward) normal $\boldsymbol{n}$ (see Fig. 1). We calculate $a = \boldsymbol{n}\cdot(\boldsymbol{s}-\boldsymbol{r})$; if $a > 0$, then the atom is inside the surface, and an extra force $\boldsymbol{F} = F\boldsymbol{n}$ is applied to it; if $a < 0$, no extra force is applied. For lipid molecules, the extra force may be averaged over the entire molecule and projected to lie in the plane of the bilayer. The closest vertex to each atom is re-calculated every few MD timesteps. Thus, an outwards force is applied to every atom inside the exclusion region, in the direction normal to the surface that delimits this region. This stage is conducted in two or more steps, by increasing the force applied from one step to the next, and attempting to attain a steady state in each case. This steady state, in which the number of atoms found within the protein surface remains approximately constant, is due to the fact that an equilibrium is established between the number of atoms exiting the region and those entering, or between the forces acting on a given atom in opposite directions. An approximate calculation, treating every atom independently, suggests that at equilibrium the number of atoms inside the surface should be inversely proportional to the exclusion force pushing them out. Hence, in order to empty the exclusion region, and at the same time minimize the perturbation of the lipid membrane, the value of the forces must be increased gradually, rather than simply extending the simulation time. Here, the insertions of KcsA and FhuA were conducted in two steps; the forces applied were $F_1 = 1$ kJ mol^{-1} $Å^{-1}$ and $F_2 = 10$ kJ mol^{-1} $Å^{-1}$ (for reference, the typical force associated with a C-C covalent bond is ca. 10^3 kJ mol^{-1} $Å^{-1}$). For these early stages, forces outside this range were shown to be either inefficient or the cause of large perturbations on the lipid bilayer (data not shown); larger forces, however, can be used effectively in subsequent stages if required (G. Patargias, personal communication). Whilst forces were applied, restraints on the z-coordinate of the lipid headgroups were used in order to prevent unnecessary bilayer distortion and separation of the bilayer leaflets. In both cases, the protein surface was positioned with respect to the bilayer so as to maximize the number of contacts between aromatic sidechains and lipid headgroups in the final system.
5. Insertion of the protein in the protein surface-shaped cavity, and re-solvation of the protein-bilayer system using a pre-equilibrated box of SPC waters (KcsA) or by a 0.1 M NaCl solution (in which the ions were placed randomly; FhuA). In both cases, a small number of additional ions are added to counteract the net charge of the protein and thus give overall electroneutrality.
6. The system was then energy minimized and finally lipid, ions and water were allowed to relax for 50 ps of MD, during which the positions of the protein atoms (except for H atoms) were restrained.

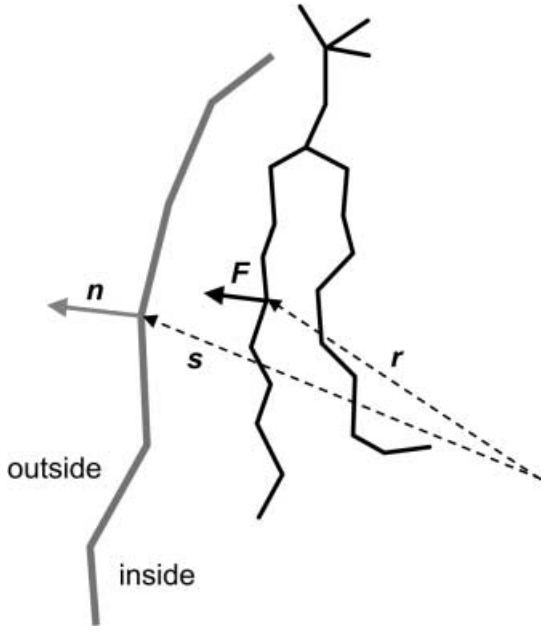

Fig. 1 Schematic diagram showing a (e.g. lipid) molecule (*black lines*) inside a protein solvent-accessible surface (*thick grey line*). The atom with vector position $\boldsymbol{r}$, within the surface, experiences a force $\boldsymbol{F}$ towards the closest surface vertex, with position vector $\boldsymbol{s}$, in the direction normal to the surface $\boldsymbol{n}$

Proteins studied

FhuA from *Escherichia coli* is an outer membrane receptor and transporter of ferrichrome, a complex produced by microorganisms to solubilize ferric iron. The overall structure of FhuA (at 2.7 Å resolution) (Ferguson et al. 1998; Locher et al. 1998) is a 22-stranded β-barrel, obstructed by a 160-amino acid N-terminal domain that folds back into the channel. KcsA (Schrempf et al. 1995) from *Streptomyces lividans* is a potassium channel, homologous to those found in, for example, the nervous systems of

animals. The crystal structure of KcsA (at 3.2 Å resolution) (Doyle et al. 1998) reveals that this integral membrane protein is made up of four identical subunits, each of which consists of two tilted transmembrane helices linked by a 30-amino acid loop-helix-loop motif, which forms the actual pore through which ions diffuse. The geometrical shape of the transmembrane regions of FhuA (Fig. 2A) and KcsA (Fig. 2B) make these proteins ideal examples of the applicability of the method presented in this paper: that of FhuA is roughly an elliptical right prism, whereas KcsA resembles a truncated cone of clover-like cross section.

Simulations

System details are summarized in Table 1. Each protein was inserted into a DMPC bilayer, initially made up of 288 DMPC lipids, that had previously been equilibrated during a 1.5 ns MD simulation (using PME), resulting in an area per lipid of 62.4 $Å^2$ [cf. experimental values of 60.0 $Å^2$ (Petrache et al. 2000b) and 59.7 $Å^2$ (Petrache et al. 1998) obtained at 303 K by ^2H NMR and X-ray diffraction, respectively]. After successful insertion into the DMPC bilayer using the protocol described above, a 1 ns equilibration simulation was carried out, during which the protein coordinates (except H atoms) were restrained. This was followed by a 1 ns unrestrained simulation. This procedure (restrained simulation and production run) was conducted once with long-range electrostatics interactions treated via cut-offs, and once using PME.

The X-ray structure of KcsA (Doyle et al. 1998) was adapted for simulations as described in Shrivastava and Sansom (2000). Two K^+ ions plus an intervening water molecule were present in the selectivity filter; the Glu71 and Asp80 residues were both ionized. In the case of FhuA, ionization states of the protein sidechains were adjusted according to the results of pK_A calculations (J.D. Faraldo-Gómez and M.S.P. Sansom, unpublished results).

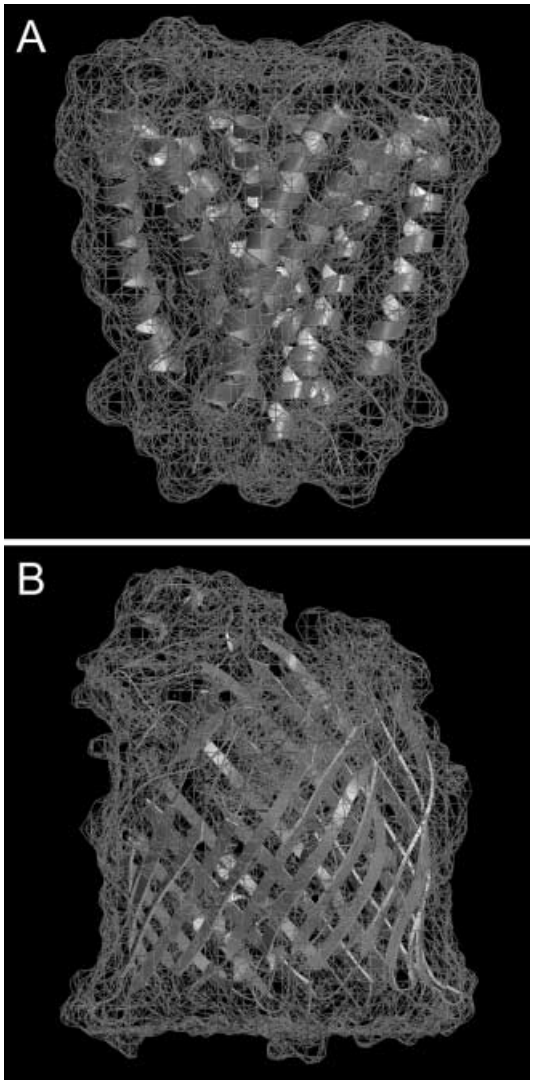

Fig. 2 The bacterial membrane proteins KcsA (**A**) and FhuA (**B**) used in this study. Solvent-accessible surfaces, generated with GRASP (Nicholls et al. 1993) and using a probe radius of 1.4 Å, are shown as a *grey mesh* superimposed upon a diagram of the protein fold. Note that on the periplasmic side of FhuA (*at the bottom*) the surface was flattened to prevent lipid molecules from becoming trapped in pockets on the surface during the optimization of the protein-lipid interface. Figs. 2, 3 and 7 were created with MOLSCRIPT (Kraulis 1991) and RASTER3D (Merritt and Bacon 1997)

Hardware and performance

Simulations were performed on an O2 SGI workstation (cavity optimization), an SGI Origin 2000 using eight parallel 195 MHz R10000 processors (FhuA-PME), or a 32-node PC cluster, using either two (KcsA-PME, KcsA-CO) or eight (FhuA-CO) parallel PIII 750 MHz processors. The performance of the PME algorithm with respect to CO was slightly worse in the two-processor (one node) simulations (about 13% slower); however, when using more than four processors, PME is very inefficient for standard ethernet-based communications, so that use of a shared-memory system is advisable.

Results

Creating a cavity

The two proteins selected for this investigation present rather different shapes, neither well approximated by a simple cylinder, which need to be embedded in a membrane; the resultant solvent-accessible surfaces (i.e. GRASP surfaces) are shown in Fig. 2. In Table 2 we provide the volumes contained by the TM regions of the two surfaces (of FhuA and KcsA). Note that, in the case of KcsA 19, more DMPC molecules were removed from the outer than from the inner leaflet, reflecting the conical shape of the protein molecule; this corresponds to a difference in cross-sectional area of ca. 1200 $Å^2$.

In Fig. 3 we illustrate the process of emptying the exclusion volume in order to optimize the cavity into which the protein is inserted. Initially a number of hydrocarbon chains are found within the protein surfaces, together with several hundred water molecules. At the end of the first stage of the insertion protocol, after 20 ps, the outwards forces applied have emptied most of the exclusion volume. However, a significant number of atoms remain within the protein surface. Finally, at $t = 25$ ps, only a few tens of atoms can be found at the surface. It should be noted that, as mentioned above, the surfaces used are solvent-accessible surfaces of the proteins. Hence, once excluded from the volume to be occupied by the protein, the conformation of the lipid molecules at the interface is complementary to the shape of the protein surface.

Table 2 Results of simulation set-up

	Upper (extracellular) leaflet		Lower leaflet	
	Volume (Å^3)[a]	Lipids removed	Volume (Å^3)[a]	Lipids removed
FhuA	27,143	32	27,478	32
KcsA	24,928	31	13,042	12

[a]Estimated volume V occupied by the solvent-accessible surface of FhuA and KcsA across the transmembrane region. Initially, lipid molecules whose headgroups are inside a cylinder of radius equal to R_{EQ} (where $V = \pi R_{EQ}^2 h$, h being the height of the volume considered, i.e. half the bilayer width) are removed, in order to preserve the density of the bilayer during the optimization of the cavity

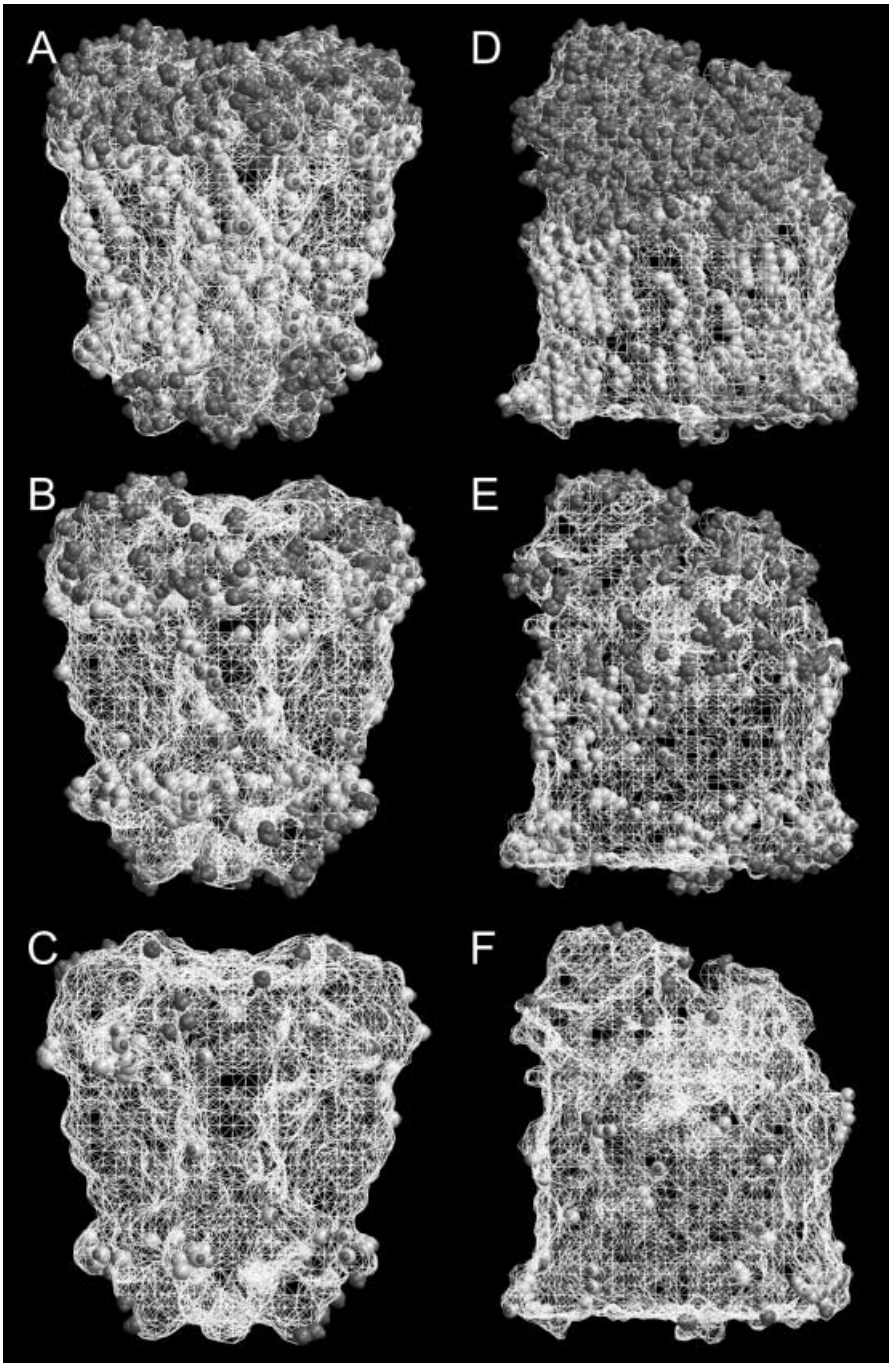

Fig. 3 Lipid (*pale grey spheres*) and water (*dark grey spheres*) molecules within the protein solvent-accessible surface (*white network*) of KcsA (**A**, **B**, **C**) and FhuA (**D**, **E**, **F**), at three stages during the cavity optimization (at **A**, **D**, $t = 0$ ps; **B**, **E**, $t = 20$ ps; and **C**, **F**, $t = 25$ ps). The directions of the outwards forces applied are shown as *small dark circles* on the atoms

The cavity optimization can be monitored in a more quantitative manner by determining the number of atoms within the molecular surface, and also the maximum depth at which they are found, as a function of simulation time (Fig. 4). In the case of FhuA, initially approximately 4400 atoms were inside the exclusion region, the deepest of these atoms being ca. 10 Å away from the protein surface. Upon the application of a force $F_1 = 1$ kJ mol^{-1} Å^{-1}, the number of atoms within the surface decreases, until a steady state was reached in about 10 ps, after which the number of atoms inside the exclusion region remained fairly constant at approximately 800. At this point, the maximum depth at which a lipid or water atom could be found was ca. 5 Å. Imposing a 10-fold greater force ($F_2 = 10$ kJ mol^{-1} Å^{-1}) for 5 ps cleared the region of lipid and water molecules almost completely. The steady state attained was then characterized by roughly 100 atoms remaining in the exclusion region to a depth of 2–3 Å. This depth might be regarded as a valid endpoint, given that solvent-accessible surfaces are used, and therefore protein and lipid atoms will not overlap significantly when inserting the protein into the cavity. A similar pattern was found in the case of KcsA. Thus we may conclude that for these forces a 20 + 5 ps simulation is sufficient to create a suitable cavity in a bilayer for a protein.

Assessing the bilayer

In order to determine the extent to which the application of the exclusion forces to lipid atoms affects the bilayer beyond the cavity-lipid interface, in Fig. 5 we plot the number of lipid atoms (of various chemical types) as a function of distance from the surface in each system. Note that even in a fully equilibrated system, with a finite simulation box, one would expect this graph to rise to a clear maximum and then decline as the edges of the box are approached. Careful comparison reveals that, as lipid atoms become expelled from inside the surface of zthe protein, a single corresponding peak of number of atoms does not appear elsewhere; rather, there is a small and distributed increase in the number of atoms at all distances away from the surface, with the most pronounced effect being at 15 Å or more. The profiles of the number density of lipid atoms along the *z*-axis also remain unchanged (data not shown), as a consequence of the use of positional restraints on the lipid headgroups in the *z*-direction. This confirms that these restraints prevent any artefactual thickening of the bilayers. Altogether, these analyses suggest that the cavity optimization method achieves a rearrangement of the lipid molecules at the interface, without grossly altering the configuration of the rest of the bilayer.

Further confirmation of this can be obtained by analysing the change in number density of the lipid atoms. The density is evaluated, using 3D grids composed of 4×4×35 Å cells, at the start and end of a period of time, and the change in density is then plotted in a 2D projection of the cells. If we apply this to two snapshots of the equilibrated DMPC bilayer (without a cavity; Fig. 6A), then there are no significant changes in density. In contrast, for FhuA (Fig. 6B, C; similar

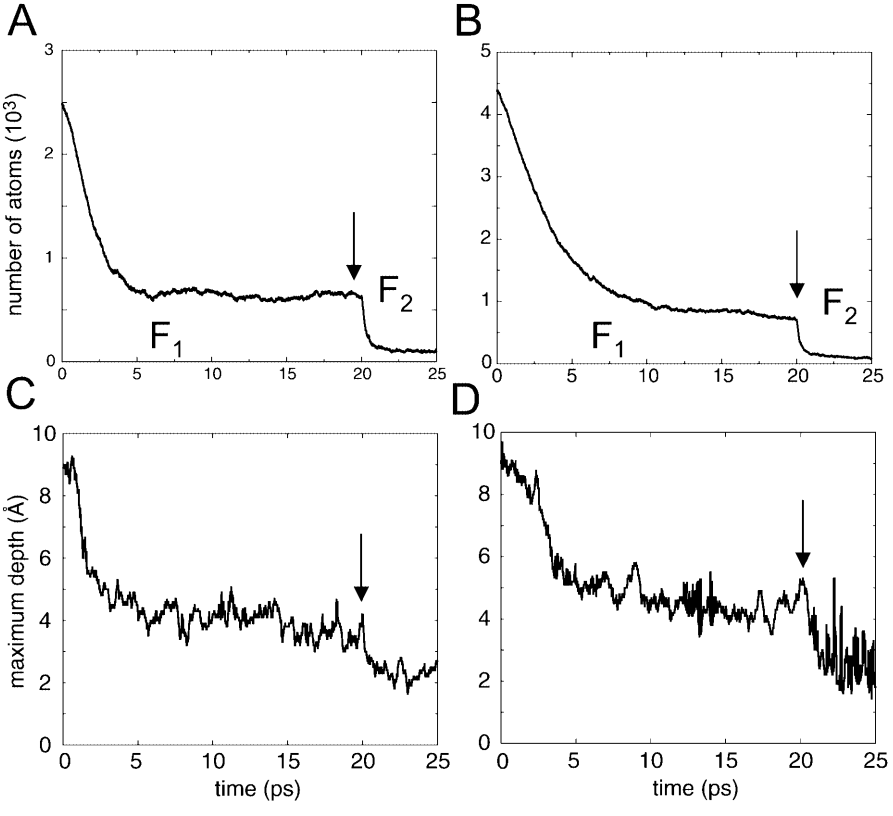

Fig. 4 Number of atoms within the solvent-accessible surface of the proteins as a function of time for **A** KcsA and **B** FhuA. Maximum depth beneath the surface at which an atom is found versus time for **C** KcsA and **D** FhuA. In all four graphs the *vertical arrow* indicates the time at which the force applied was increased from $F_1 = 1$ kJ mol^{-1} Å^{-1} to $F_2 = 10$ kJ mol^{-1} Å^{-1}

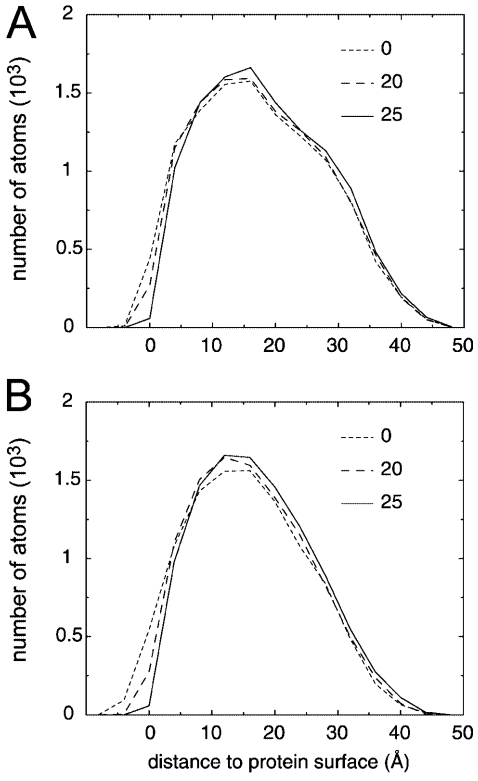

Fig. 5 Number of lipid atoms versus distance from the protein solvent-accessible surface for three snapshots [$t = 0$ (*grey*, *dotted lines*), $t = 20$ ps (*black*, *broken lines*) and $t = 25$ ps (*black*, *solid lines*)] during the cavity-optimization process for **A** KcsA and **B** FhuA. Distances were calculated using the coordinates of lipid atoms and of their respective closest surface vertices (i.e. $|\boldsymbol{s}-\boldsymbol{r}|$, as defined in Fig. 1)

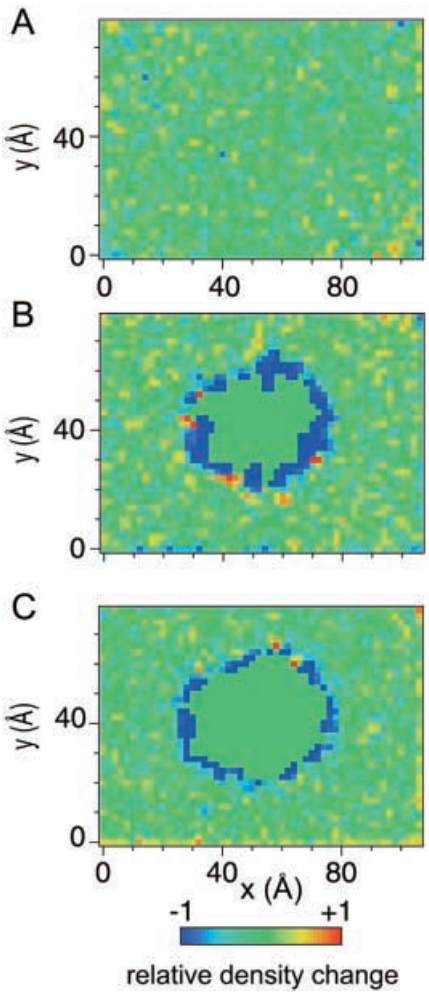

Fig. 6 Two-dimensional maps of the relative changes in lipid atom density for **A** an equilibrated DMPC bilayer over a 20 ps period; **B** the FhuA system showing the density change from $t = 0$ to 20 ps; and **C** the same system for $t = 20$ to 25 ps. The number density for each snapshot v is calculated using two 3D grids of 4×4×35 Å^3 cells, shifted in the x and y direction with respect to each other by 2 Å (hence the cells overlap by 50%) in order to improve the sampling; a large bin size in the z direction is used so as to average out inhomogeneities in this direction. The relative density change f is calculated as $f = \{(v_{\mathrm{FINAL}} - v_{\mathrm{INITIAL}})/max(v_{\mathrm{FINAL}}, v_{\mathrm{INITIAL}})\}$. Thus, f is dimensionless and is positive (*red*) if the relative density increases over the period of time, and negative (*blue*) if it decreases

results are seen for KcsA – data not shown) this analysis clearly reveals a 5-Å thick annulus of negative relative number density (i.e. loss of atoms) around the surface of the cavity, as can be anticipated. Equally clearly, there is no significant systematic increase in relative number density around the negative annulus; instead, there are small scattered increases in density, although these are not all pronounced and are almost absent for the $t = 20$ versus 25 ps comparison (Fig. 6C).

Simulation behaviour

The configuration of the FhuA system generated as a result of this procedure is shown in Fig. 7. It appears that the DMPC molecules are efficiently packed around the protein surfaces. A similar conclusion is drawn from inspecting images of the KcsA/DMPC system (not shown).

To evaluate the "quality" of these systems we have performed two sets of 1-ns simulations for each protein. In these simulations we have explored the use of cut-offs versus PME to treat long-range electrostatic interactions. There has been some discussion (Tieleman 1998; Tobias 2001; Tobias et al. 1997) of the relative merits of these approaches for membrane simulations, and we wished to examine their influence on the optimization of simulations of membrane-protein systems.

In order to evaluate the "stability" of a simulation, we analysed the extent to which the structure of a protein drifts away from that in the crystal structure, in terms of the Cα root mean square deviation (RMSD) relative to the initial structure (Fig. 8). Typically, for a stable simulation one would expect an initial rise in Cα RMSD over the first 100 ps or so (due to the relaxation of the protein in the non-crystallographic environment), followed by a constant RMSD with respect to time (indicative of no further net drift in structure) for the remainder of the simulation; this is roughly the case

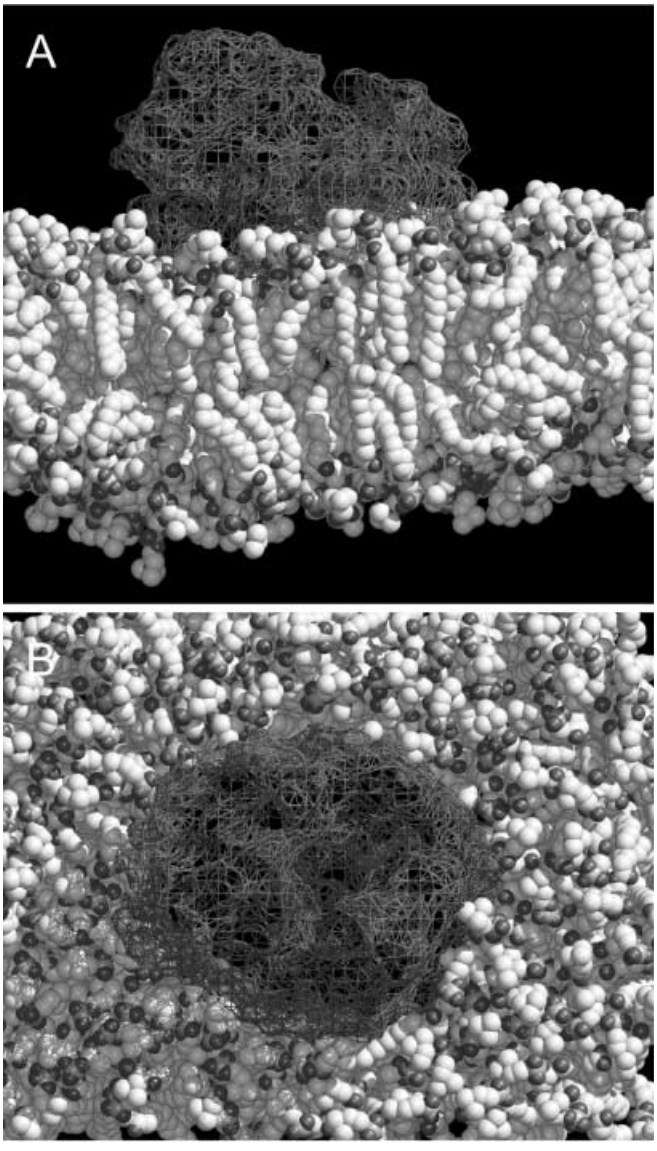

Fig. 7 The FhuA surface-lipid bilayer system at the end of the cavity-optimization process, viewed perpendicular to (**A**) and along (**B**) the bilayer normal. Lipid molecules are shown as van der Waals surfaces, protein as a GRASP mesh surface

for both the KcsA and the FhuA simulations in this study. For both simulations the Cα RMSD is slightly increased if the long-range interactions are treated with a cut-off than if PME is used; this difference is less if only the core secondary structure elements are considered, suggesting that the structural drift occurs in the non-membranous loops of the two proteins. The Cα RMSD for KcsA in these simulations is the same as that seen in earlier simulations (Shrivastava and Sansom 2000) of KcsA/POPC using cut-offs, and in simulations of Kir6.2, a K-channel model derived from the KcsA structure (Capener et al. 2000). It is tempting to conclude that the structural drift is lower for FhuA than for KcsA, which could reflect either the different folds and/or the different resolutions of the two structures.

Discussion

Insertion methodology

The simulations presented in this paper suggest that a cavity-creation-based methodology for insertion of complex membrane proteins into lipid bilayers is feasible, and thus is an alternative to methods based on the assembly of a bilayer around a membrane protein (Petrache et al. 2000a; Roux and Woolf 1996; Tang et al. 1999; Woolf and Roux 1994, 1996). It might be interesting to compare simulations based on these two methods. The results presented in this study suggest that the current method generates a relatively unperturbed bilayer; provided this criterion is satisfied, one might not anticipate profound differences between systems generated by different methods. A cavity insertion method might have some advantages for complex membrane systems where generation of the initial bilayer has already proved to be quite complex and/or costly computationally (Lins and Straatsma 2001). Such methods may also be helpful for the semi-automated set-up of simulations of a wide range of membrane proteins that will be necessary as the rate of experimental structure determination increases. However, some methodological refinements may be necessary to take into account some of the subtleties of lipid-protein interactions that are now beginning to be revealed by X-ray crystallography (Fyfe et al. 2001), such as the localization of acyl chains in grooves in the protein surface, owing to hydrophobic interactions with apolar residues and backbone atoms.

Simulation optimization

Insertion into a bilayer is only one component of optimizing a simulation of a membrane protein; the results presented here indicate that the treatment of long-range electrostatics can have an effect on overall protein dynamics during such simulations. At first glance one might conclude that the results demonstrate that, for the proteins studied here, PME is more appropriate, since the inclusion of all the electrostatic interactions yielded lower drift from the X-ray conformation, particularly for the loop regions. However, studies on simple peptides have suggested artefactual stabilization of folded (α-helical) conformations (i.e. low RMSD) by the use of particle-mesh methods in small systems (Weber et al. 2000), so care has to be taken in considering suitably large system sizes in order to prevent similar artefacts. Furthermore, although the overall Cα RMSD values for KcsA and FhuA were lower in simulations using PME than in those using cut-offs, this pattern is not always repeated. For example, in the case of simulations of a Kir6.2 channel model we have seen somewhat smaller effects than observed here (Capener and Sansom 2002). Comparing two related proteins, GlpF (Fu et al. 2000) and Aqp1 (Murata et al. 2000; Ren et al. 2001), in the

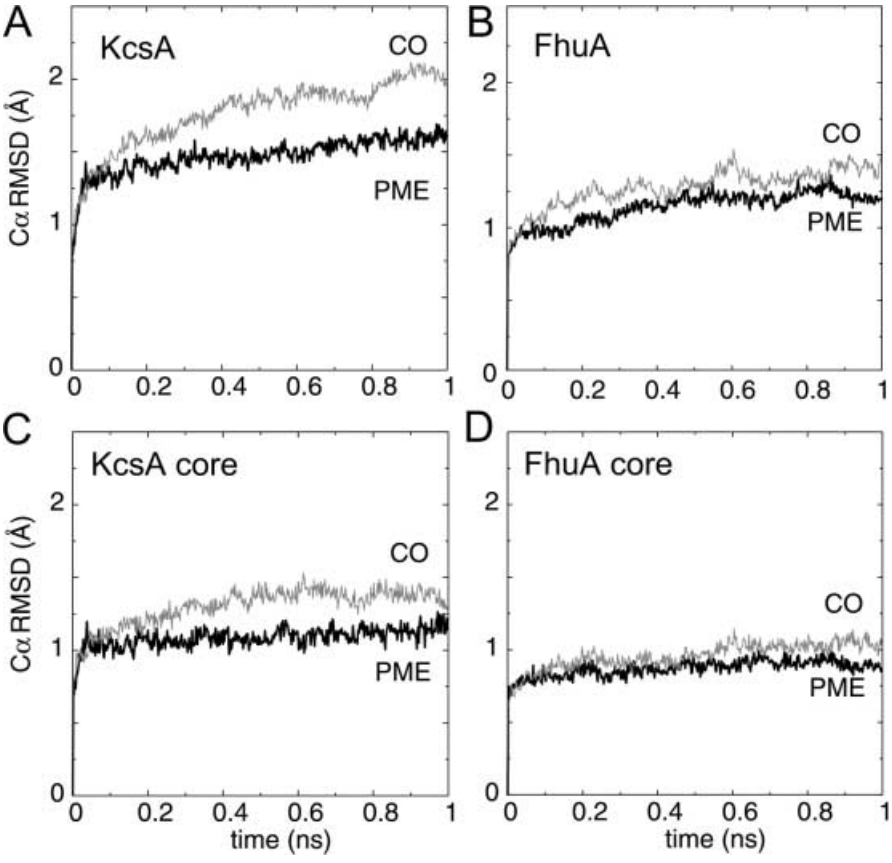

Fig. 8A–D Root mean square deviation (RMSD) relative to the X-ray structures as a function of time for **A**, **C** KcsA and **B**, **D** FhuA. In each graph, RMSD curves are shown for the cut-off (CO) and PME simulations. In **A**, **B**, RMSD values for all Cα atoms of the protein are shown; in **C**, **D**, RMSD values were calculated for just the core secondary structure Cα atoms, i.e. the Cα atoms of the M1-helix, P-helix, selectivity filter and M2-helix regions of KcsA or the Cα atoms of the β-strands of FhuA

former case there is a significant difference between PME and cut-off simulations, whereas for the latter there is little, if any, difference (Patargias, Law, Sansom, unpublished observations). At the moment, a conservative conclusion would be to say that in some cases the use of PME rather than cut-off results in a lower Cα RMSD, but that each new system may merit further investigation.

Future challenges

In this paper we have analysed a cavity optimization method for generating an initial configuration for simulation of integral membrane proteins, i.e. those membrane proteins that span a lipid bilayer. A somewhat more complicated task is presented by those peptides or proteins that are associated with the surface of a lipid bilayer. Simulations to date (Bernèche et al. 1998; Nina et al. 2000; Sankararamakrishnan and Weinstein 2000; Tieleman et al. 1999b) have to some extent relied on manually guided methods to generate initial configurations. More automated approaches to peptide-bilayer interactions (Tieleman et al. 2001) require extended simulation times and, for lipid bilayers rather than membrane mimetics such as an octane slab, there remain worries over convergence. It will be important to explore the applicability of cavity-based approaches to such systems, guided by relevant experimental data.

Acknowledgements This work was supported by grants from The British Council-Fundación La Caixa, EPSRC and the Wellcome Trust. The Oxford Supercomputing Centre provided computational facilities. We are most grateful to L.R. Forrest, D.P. Tieleman, I.H. Shrivastava and P.J. Bond for valuable discussions.

References

Berendsen HJC, Postma JPM, Gunsteren WF van, DiNola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690

Berendsen HJC, Spoel D van der, Drunen R van (1995) GROMACS: a message-passing parallel molecular dynamics implementation. Comput Phys Comm 95:43–56

Berger O, Edholm O, Jahnig F (1997) Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure and constant temperature. Biophys J 72:2002–2013

Bernèche S, Nina M, Roux B (1998) Molecular dynamics simulation of melittin in a dimyristoylphosphatidylcholine bilayer membrane. Biophys J 75:1603–1618

Capener CE, Sansom MSP (2002) MD Simulations of a K channel model – sensitivity to changes in ions, waters and membrane environment. J Phys Chem B (in press)

Capener CE, Shrivastava IH, Ranatunga KM, Forrest LR, Smith GR, Sansom MSP (2000) Homology modelling and molecular dynamics simulation studies of an inward rectifier potassium channel. Biophys J 78:2929–2942

Chiu SW, Jacobson KA, Scott HL (2001) Combined Monte Carlo and molecular dynamics simulation of hydrated lipid-cholesterol lipid bilayers at low cholesterol concentration. Biophys J 80:1104–1114

Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. J Chem Phys 98:10089–10092

Doyle DA, Cabral JM, Pfuetzner RA, Kuo A, Gulbis JM, Cohen SL, Cahit BT, MacKinnon R (1998) The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. Science 280:69–77

Feller SE, Pastor RW (1999) Constant surface tension simulations of lipid bilayers: the sensitivity of surface areas and compressibilities. J Chem Phys 111:1281–1287

Ferguson AD, Hofmann E, Coulton JW, Diederichs K, Welte W (1998) Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. Science 282:2215–2220

Forrest LR, Sansom MSP (2000) Membrane simulations: bigger and better? Curr Opin Struct Biol 10:174–181

Forrest LR, Tieleman DP, Sansom MSP (1999) Defining the transmembrane helix of M2 protein from influenza A by molecular dynamics simulations in a lipid bilayer. Biophys J 76:1886–1896

Forstner MB, Käs J, Martin D (2000) Single lipid diffusion in Langmuir monolayers. Langmuir 17:567–570

Fu D, Libson A, Miercke LJW, Weitzman C, Nollert P, Krucinski J, Stroud RM (2000) Structure of a glycerol-conducting channel and the basis of its selectivity. Science 290:481–486

Fyfe PK, McAuley KE, Roszak AW, Isaacs NW, Codgell RJ, Jones MR (2001) Probing the interface between membrane proteins and membrane lipids by X-ray crystallography. Trends Biochem Sci 26:106–112

Gennis RB (1989) Biomembranes: molecular structure and function. Springer, Berlin Heidelberg New York

Govaerts C, Blanpain C, Deupi X, Ballet S, Ballesteros JA, Wodak SJ, Vassart G, Pardo L, Parmentier M (2001) The TXP motif in the second transmembrane helix of CCR5 – a structural determinant of chemokine-induced activation. J Biol Chem 276:13217–13225

Guidoni L, Torre V, Carloni P (1999) Potassium and sodium binding in the outer mouth of the K^+ channel. Biochemistry 38:8599–8604

Hermans J, Berendsen HJC, Gunsteren WF van, Postma JPM (1984) A consistent empirical potential for water-protein interactions. Biopolymers 23:1513–1518

Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: a linear constraint solver for molecular simulations. J Comput Chem 18:1463–1472

Kraulis PJ (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J Appl Crystallogr 24:946–950

Lindahl E, Edholm O (2000) Mesoscopic undulations and thickness fluctuations in lipid bilayers from molecular dynamics simulations. Biophys J 79:426–433

Lins RD, Straatsma TP (2001) Computer simulation of the rough lipopolysaccharide membrane of *Pseudomonas aeruginosa*. Biophys J 81:1037–1046

Locher KP, Rees B, Koebnik R, Mitschler A, Moulinier L, Rosenbusch J, Moras D (1998) Transmembrane signalling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes. Cell 95:771–778

Merritt EA, Bacon DJ (1997) Raster3D photorealistic molecular graphics. Methods Enzymol 277:505–524

Murata K, Mitsuoaka K, Hirai T, Walz T, Agre P, Heymann JB, Engel A, Fujiyoshi Y (2000) Structural determinants of water permeation through aquaporin-1. Nature 407:599–605

Nicholls A, Bharadwaj R, Honig B (1993) GRASP: graphical representation and analysis of surface properties. Biophys J 64:166–170

Nina M, Bernèche S, Roux B (2000) Anchoring of a monotopic membrane protein: the binding of prostaglandin H2 synthase-1 to the surface of a phospholipid bilayer. Eur Biophys J 29:439–454

Pastor RW, Feller SE (1996) Time scales of lipid dynamics and molecular dynamics. In: Merz KM (ed) Biological membranes: a molecular perspective from computation and experiment. Birkhäuser, Boston, p 587

Petrache HI, Tristram-Nagle S, Nagle JF (1998) Fluid phase structure of EPC and DMPC bilayer. Chem Phys Lipids 95:83–94

Petrache HI, Grossfield A, MacKenzie KR, Engelman DM, Woolf TB (2000a) Modulation of glycophorin A transmembrane helix interactions by lipid bilayers: molecular dynamics calculations. J Mol Biol 302:727–746

Petrache HI, Dodd SW, Brown MF (2000b) Area per lipid and acyl length distributions in fluid phosphatidylcholines determined by ^2H NMR spectroscopy. Biophys J 79:3172–3192

Randa HS, Forrest LR, Voth GA, Sansom MSP (1999) Molecular dynamics of synthetic leucine-serine ion channels in a phospholipid membrane. Biophys J 77:2400–2410

Ren G, Reddy VS, Cheng A, Melnyk P, Mitra AK (2001) Visualization of a water-selective pore by electron crystallography in vitreous ice. Proc Natl Acad Sci USA 98:1398–1403

Roux B, Woolf TB (1996) Molecular dynamics of Pf1 coat protein in a phospholipid bilayer. In: Merz KM (ed) Biological membranes: a molecular perspective from computation and experiment. Birkhäuser, Boston, pp 555–587

Roux B, Bernèche S, Im W (2000) Ion channels, permeation and electrostatics: insight into the function of KcsA. Biochemistry 39:13295–13306

Sankararamakrishnan R, Weinstein H (2000) Molecular dynamics simulations predict a tilted orientation for the helical region of dynorphin A (1–17) in dimyristoylphosphatidylcholine bilayers. Biophys J 79:2331–2344

Sanner MF, Olson AJ, Spehner JC (1996) Reduced surface: an efficient way to compute molecular surfaces. Biopolymers 38:305–320

Sansom MSP, Shrivastava IH, Ranatunga KM, Smith GR (2000) Simulations of ion channels: watching ions and water move. Trends Biochem Sci 25:368–374

Schneider MJ, Feller SE (2001) Molecular dynamics simulations of a phospholipid-detergent mixture. J Phys Chem B 105:1331–1337

Schrempf H, Schmidt O, Kummerlein R, Hinnah S, Muller D, Betzler M, Steinkamp T, Wagner R (1995) A prokaryotic potassium-ion channel with 2 predicted transmembrane segments from *Steptomyces lividans*. EMBO J 14:5170–5178

Shen L, Bassolino D, Stouch T (1997) Transmembrane helix structure, dynamics, and interactions: multi-nanosecond molecular dynamics simulations. Biophys J 73:3–20

Shrivastava IH, Sansom MSP (2000) Simulations of ion permeation through a potassium channel: molecular dynamics of KcsA in a phospholipid bilayer. Biophys J 78:557–570

Smart OS, Goodfellow JM, Wallace BA (1993) The pore dimensions of gramicidin A. Biophys J 65:2455–2460

Smart OS, Neduvelil JG, Wang X, Wallace BA, Sansom MSP (1996) HOLE: a program for the analysis of the pore dimensions of ion channel structural models. J Mol Graphics 14:354–360

Tang YZ, Chen WZ, Wang CX, Shi YY (1999) Constructing the suitable initial configuration of the membrane-protein system in molecular dynamics simulations. Eur Biophys J 28:478–488

Terstappen GC, Reggiani A (2001) In silico research in drug discovery. Trends Pharmacol Sci 22:23–26

Tieleman DP (1998) Theoretical studies of membrane models. Biophysical Chemistry, Rijksuniversiteit Groningen, Groningen, p 192

Tieleman DP, Berendsen HJC (1998) A molecular dynamics study of the pores formed by *Escherichia coli* OmpF porin in a fully hydrated palmitoyloleoylphosphatidylcholine bilayer. Biophys J 74:2786–2801

Tieleman DP, Marrink SJ, Berendsen HJC (1997) A computer perspective of membranes: molecular dynamics studies of lipid bilayer systems. Biochim Biophys Acta 1331:235–270

Tieleman DP, Berendsen HJC, Sansom MSP (1999a) An alamethicin channel in a lipid bilayer: molecular dynamics simulations. Biophys J 76:1757–1769

Tieleman DP, Berendsen HJC, Sansom MSP (1999b) Surface binding of alamethicin stabilises its helical structure: molecular dynamics simulations. Biophys J 76:3186–3191

Tieleman DP, Berendsen HJC, Sansom MSP (2001) Voltage-dependent insertion of alamethicin at phospholipid/water and octane/water interfaces. Biophys J 80:331–346

Tobias DJ (2001) Electrostatics calculations: recent methodological advances and applications to membranes. Curr Opin Struct Biol 11:253–261

Tobias DJ, Tu KC, Klein ML (1997) Atomic-scale molecular dynamics simulations of lipid membranes. Curr Opin Coll Interface Sci 2:15–26

Venable RM, Zhang Y, Hardy BJ, Pastor RW (1993) Molecular dynamics simulations of a lipid bilayer and of hexadecane: an investigation of membrane fluidity. Science 262:223–226

Wallin E, Heijne G von (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archean, and eukaryotic organisms. Protein Sci 7:1029–1038

Weber W, Hunenberger PH, McCammon JA (2000) Molecular dynamics simulations of a polyalanine octapeptide under Ewald boundary conditions: influence of artificial periodicity on peptide conformation. J Phys Chem B 104:3668–3675

Woolf TB, Roux B (1994) Molecular dynamics simulation of the gramicidin channel in a phospholipid bilayer. Proc Natl Acad Sci USA 91:11631–11635

Woolf TB, Roux B (1996) Structure, energetics, and dynamics of lipid-protein interactions: a molecular dynamics study of the gramicidin-A channel in a DMPC bilayer. Proteins Struct Funct Genet 24:92–114